4-13-2017

Background Subtraction in Video Surveillance

Shanpreet Kaur
University of Windsor

Background subtraction in Video Surveillance

by

Shanpreet Kaur

A Thesis
Submitted to the Faculty of Graduate Studies
Through Computer Science
In Partial Fulfilment of the Requirements for
The Degree of Master of Science at the
University of Windsor

Windsor, Ontario, Canada

2016

Background Subtraction in Video Surveillance

by

Shanpreet Kaur

APPROVED BY:

Dr. Faouzi Ghrib
Department of Civil and Environmental Engineering

Dr. Imran Ahmad
School of Computer Science

Dr. Boubakeur Boufama, Advisor
School of Computer Science

December 19, 2016

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication. I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree of any other University or Institution.

ABSTRACT

The aim of thesis is the real-time detection of moving and unconstrained surveillance environments monitored with static cameras. This is achieved based on the results provided by background subtraction. For this task, Gaussian Mixture Models (GMMs) and Kernel density estimation (KDE) are used. A thorough review of state-of-the-art formulations for the use of GMMs and KDE in the task of background subtraction reveals some further development opportunities, which are tackled in a novel GMM-based approach incorporating a variance controlling scheme. The proposed approach method is for parametric and non-parametric and gives us the better method for background subtraction, with more accuracy and easier parametrization of the models, for different environments. It also converges to more accurate models of the scenes.

The detection of moving objects is achieved by using the results of background subtraction. For the detection of new static objects, two background models, learning at different rates, are used. This allows for a multi-class pixel classification, which follows the temporality of the changes detected by means of background subtraction.

In a first approach, the subtraction of background models is done for parametric model and their results are shown. The second approach is for non-parametric models, where background subtraction is done using KDE non-parametric model.

Furthermore, we have done some video engineering, where the background subtraction algorithm was employed so that, the background from one video and the foreground from another video are merged to form a new video. By doing this way, we can also do more complex video engineering with multiple videos.

Finally, the results provided by region analysis can be used to improve the quality of the background models, therefore, considerably improving the detection results.

DEDICATION

To the almighty God, my parents Daljinder Singh and Jagdish Kaur, Brother Arshwinder Singh.

ACKNOWLEDGEMENT

I would like to take this opportunity to thank my supervisor Dr. Boubakeur Boufama for his encouragement and support in presenting this thesis work. My ultimate gratitude goes to him for contributing his suggestions and ideas during my research. His insightful feedback and instructions made it possible for me to accomplish this work.

I would like to acknowledge my thesis committee members Dr. Imran Ahmad and Dr. Faouzi Ghrib whose suggestions and recommendations greatly improved the quality of this work. I would like to thank them for spending their valuable time providing feedback about thesis throughout my proposal and defense.

My special thanks goes to my parents and my brother for their patience and love they provided to me during all times. I express my deep appreciation to my all friends for their motivation and moral support they provided during all stages of my thesis work. I would like to thank students of Computer Science and all other departments who participated in the usability study of my thesis. I am deeply grateful for the time and effort they spent on the test.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS/SYMBOLS

KDE - Kernel Density Estimation

GMM - Gaussian Mixture Model

PTZ - Pan Tilt Zoom

IVS - Intelligent Video Surveillance

VA - Video Analytic

CHAPTER 1

INTRODUCTION

Computer vision, image processing and pattern recognition are vast fields of research concerning the automatic analysis of images and image sequences, with a broad spectrum of applications such as remote sensing, medical diagnosis, human-computer interaction or video compression, to mention only a few of them. Benefitting from the advances in those fields, robotized video-based reconnaissance has emerged as a claim inquire about point which has picked up a ton of consideration in the late years, because of the expanding dangers to the security openly places, for example, railroad stations or airplane terminals. The point is to help human administrators in checking Closed Circuit Television (CCTV) camera systems, by alarming them on deviation from the typical conduct saw in the region under observation. This gives the fundamental advantage that an administrator may screen a bigger measure of cameras by concentrating his regard for the basic focuses in space and time, while the framework expects the dreary undertaking of checking regions where non-intriguing occasions are going on. Moreover, the learning gained by method for programmed video investigation strategies can be utilized as a part of request to help video administrators and legitimate experts in the recovery of confirmation verifications from recorded video information, to regulate huge zone video arranges in assignments, for example, panning and zooming all through Pan-Tilt-Zoom (PTZ) cameras, and notwithstanding for less specialized issues as securing the protection of people in broad daylight places.

Video surveillance systems have experienced a rapid development in the last decades, especially after the attacks on the 11th of September 2001 in New York, 11th of March 2004 in Madrid and 7th and 21st of July 2005 in London, leading them to become a part of our daily life. But the use of video surveillance systems is not restricted to safety and security applications. Nowadays, video surveillance systems are also being deployed at department stores in order to provide advertising assessment and quality of service, on highways for traffic monitoring purposes, and even on houses for elderly people to assist them in a non-invasive manner. This success has been supported by the decaying prices in the sensor industry, which is able to provide higher quality cameras of ever smaller sizes at low prices. Moreover, the introduction of wireless networks has connoted a drastic reduction in the deployment costs. With the transition to IP camera networks, large camera networks can be both local and remotely controlled.

The quick development of video reconnaissance frameworks brings about an expanding number of video sustains which ought to be checked and put away in a control room. These outcomes in a constantly developing workload for CCTV administrators, who are overpowered by the gigantic arrangements of cameras. To ease this issue, programmed video examination procedures go for comprehension activities and human practices in video successions with a specific end goal to caution CCTV administrators upon the event of debilitating circumstances. This situation relates to the proactive side of wrongdoing anticipation. Besides that, video surveillance systems can also be used for crime investigation and offenders' prosecution. Video indexing and summarization can be used in order to effectively accomplish this last task. Furthermore, automated video surveillance systems have given raise to the paradigm of bringing intelligence to the edge of the network. This allows for the design of distributed surveillance networks, which require a lower bandwidth for the transmission of the captured information.

Nevertheless, as video surveillance systems have become ubiquitous, some aspects of the deployed systems have been questioned. One of the aspects is the effectiveness regarding crime prevention. Another is the need of protecting the privacy and security of personal information, which has gained increasing attention in the recent years. The Telegraph claimed that an individual will appear on average on 300 CCTV cameras during a day [Gray, 2008].

All of these aspects together have attracted the attention of both the academy and the industry, and is expected to continue growing in the next years. A recent report of Homeland Security Research Corporation [HSRC, 2013] estimates the revenue of the global Intelligent Video Surveillance (IVS) & Video Analytics (VA) industry as $13.5 billion in 2012, and predicts a rapid growth until 2020, where it is expected to reach $39 billion.

The specialized origination and sending of mechanized video-based observation frameworks include various key issues to be tended to. The most minimal level of the framework configuration concerns equipment issues, including video gaining (cameras), stockpiling gadgets and systems. At this level, choices are taken like system topology and correspondence conventions. Upon this level, the data accumulated by the cameras is investigated by method for picture and video handling systems, to separate valuable data out of the video successions. This is the level giving the semantic abilities of the framework. Finally, at the top level, the extracted information is presented to the user and eventually stored in a database for further usage. At this level, considerations on the ergonomics of the system as a whole and human-computer interaction should be taken into account. Obviously, decisions made at the different levels of design might affect the decisions to be made at the other levels; even more in the case of bringing intelligence to the network. The main focus of this thesis is set on the video processing and understanding chain.

1.1 Video-Based Surveillance Systems

Automated video-based surveillance systems, in this thesis referred to as surveillance systems for brevity (otherwise explicitly indicated), rely on the automatic detection of events of interest by means of several analysis techniques mainly stemming from the fields of computer vision, image processing and pattern recognition. Detecting events of interest is an application dependent task and can be approached in very different manners. Nevertheless, there is a common number of steps that a general surveillance system usually goes through, namely, object detection, object association, commonly referred to as tracking, and scene understanding, often accomplished by the less ambitious task of event detection. In order to successfully accomplish these tasks, the cameras have to be calibrated with respect to an extrinsic Cartesian reference space, therefore allowing for a measurement of the size and position of the detected objects. These main building blocks of an automated video-based surveillance system are depicted in Figure 1.1 and briefly introduced in the following subsections.

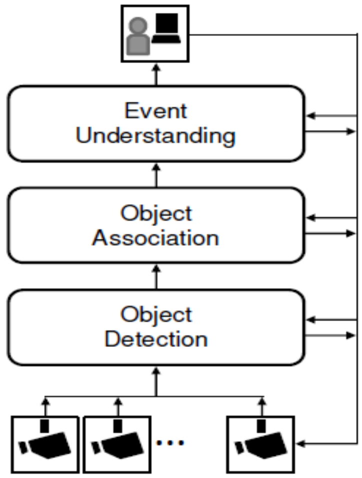


Fig 1.1 General video-based surveillance system [7].

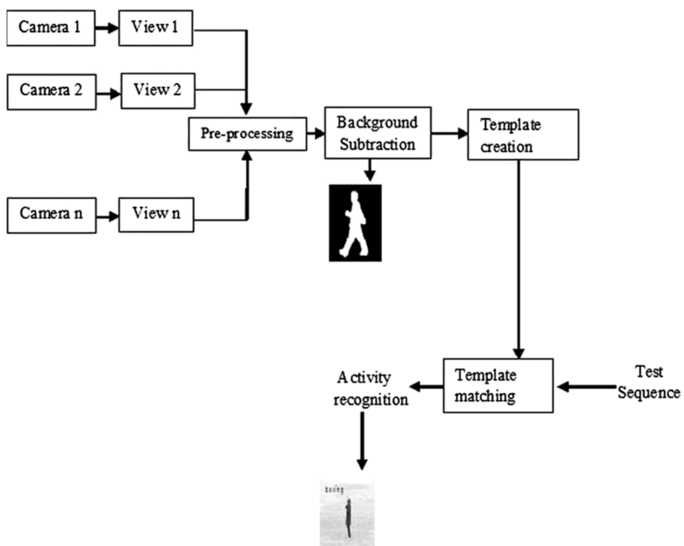


Fig 1.2 General video-based surveillance system with multiple cameras [7].

1.1.1 Object Detection and Classification

Generic object recognition, also known as category-level object recognition, is considered to be one of the most challenging visual tasks in computer vision [91]. Given any instance of a particular general class as, e.g., ’person’, ’car’ or ’bicycle’, the task is to correctly localize and classify it by means of visual features. A thorough pursuit over all protest models and picture areas can be excessively tedious for some computer vision applications. To decrease the multifaceted nature of the issue, reconnaissance frameworks for the most part partition the issue into two stages: to start with, the objects of intrigue are recognized and, second, the distinguished articles are ordered. Objects of intrigue are normally characterized as those articles presenting some sort of progress in the watched scene and are by and large related to moving items.

Object detection can be approached by means of three different techniques: temporal differencing, background subtraction, and optical flow. These three techniques give's a low-level pixel order. To assemble objects, pixels are then grouped taking care of this arrangement and their spatial setup. Worldly differencing depends on figuring the distinction of back to back video outlines at each pixel position and characterizing as changed pixels those which outright contrast surpasses a given limit. Brief differencing is exceedingly versatile to element situations and low requesting in computational terms, however it neglects to separate the entire arrangement of pixels comparing to the articles in movement. Early works based on temporal differencing can be found in [92] and references therein. Background subtraction is the most commonly used approach in setups with static cameras. It consist in using a model of the scene background in order to detect foreground objects by differencing incoming frames with the model. Background subtraction is mostly fast and has low computational demands. However, it can be sensitive to sudden illumination changes and small camera motions as, e.g., vibrations. A good introduction to background subtraction, including the main issues that a background subtraction approach has to deal with, can be found in [18]. Optical flow is an estimation used to determine corresponding points between two images. Optical flow based methods can be used to detect independently moving objects even in the presence of camera motion. Nevertheless, even in their most efficient implementations, they are highly demanding in computational terms. Furthermore, depending on the smoothness constraint, the corresponding points in the considered frames might not be allowed to be more than a few pixels away, therefore, being constrained the speed of movement of objects and camera. A good introduction to the topic of optical flow computation can be found in [12]. An overview of state-of-the-art approaches and their respective performance can be consulted on-line in the Middlebury dataset website2 [4].

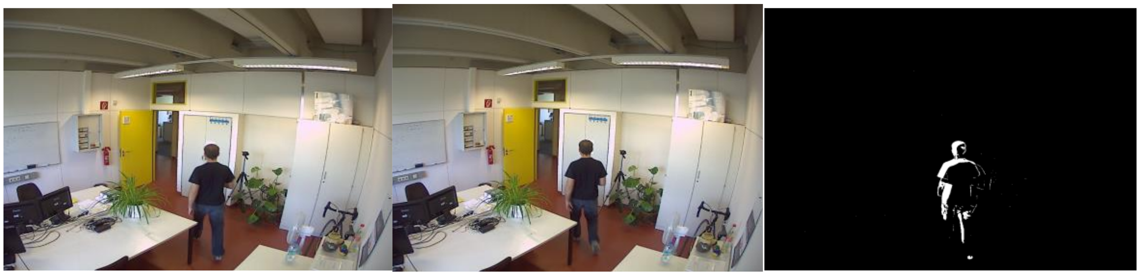


Fig 1.3**:** Temporal differencing. From left to right: first image of a pair containing one Moving person, second image of the same image pair, and difference mask.

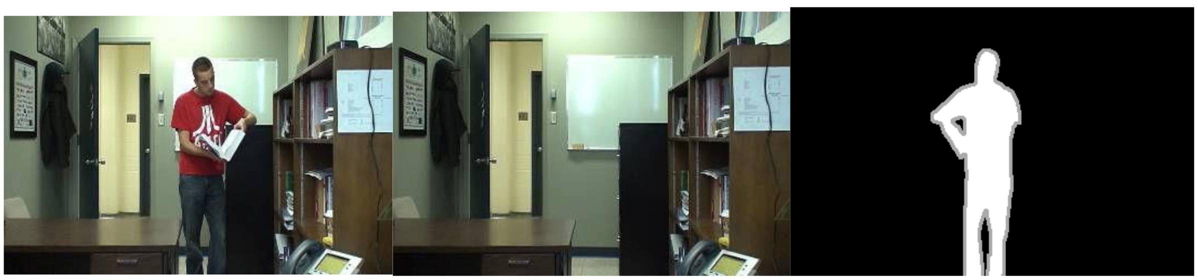


Fig 1.4**:** Background subtraction example for two frames of the sequence 'office' Background of the scene, and ground-truth foreground mask (source, www.changedetection.net).

1.1.2 Object Tracking/ motion tracking

Object tracking is the task of setting up correspondences between the distinguished protests over the casings of a video sequence. With a specific end goal to achieve this undertaking, a model for the articles and the movement they display is utilized. Ordinary question models are focuses, primitive geometric shapes, as, e.g., ovals and rectangles, outlines, explained shape models and skeletons. Contingent upon the chose question show utilized, the movement model can be delimited. For example, if an object is represented by a point, then, only a translational model can be used, whereas in the case of more elaborated object models as, e.g., silhouettes, parametric and non-parametric motion models can be used. Depending on the application domain, assumptions are made in order to constrain the tracking problem. In the surveillance domain, point-based tracking models are a popular choice to solve the tracking problem. Thereby, Kalman [93] and Particle Filters [94] are commonly state estimation methods used for computing the cost of a given object association. An excellent introduction into the tracking topic and important related issues including the use of appropriate image features, selection of motion models, and detection of objects, can be found in [95].

The object detection, can be done in many ways. How to do it depends on data available and whether the object is in motion or not. For objects at rest some prior knowledge regarding the type of objects must be known. This can be a single sample image of the object to track. Detecting moving objects in an image sequence does not need prior knowledge but needs multiple consecutive images. Two common methods for detecting moving objects are [8]:

Background subtraction

Background Subtraction is a widely-used approach for detecting moving objects from static cameras. The fundamental logic is detecting objects from a difference between the current frame and reference frame, called background image. The principle is that if a reference background image is known, that image can be compared with the frame in which objects are to be detected. The regions that are different contain moving objects.

Optical flow

By calculating the flow field of pixels in successive frames it is possible to detect objects. Clusters of pixels moving together are likely to be part of the same object.

When the location of the object to be tracked is known some features must be extracted and recorded to make it possible to find the same object in new frames. Good segmentation from the background ensures that only features that actually belong to the object of interest are recorded. The problem is thus, given an area containing an object, to determine which pixels belong to the object and which belong to the background. In some cases a pixel-wise segmentation is not needed, but if too much background gets incorporated in the object model the noise will make it very hard to keep track of the target.

Objects can be represented in multiple ways, as a centroid point, multiple points, primitive geometric shapes or object contours and silhouettes. These can be combined to get a good representation of the object that is to be tracked. Good features to track are things that continue looking the same even if the scale changes or the object rotates out of plane. Examples of that kind of features are corners and edges. Another possible representation of the object is the color histogram of the object area.

1.1.3 Background subtraction

In the most fundamental sense background subtraction is just what the name concludes, the total contrast between a reference picture (the background) and a picture of interest. At picture positions where the distinction is more prominent than some edge the position is classified not having a place with the foundation, i.e. named a forefront pixel. Most present-day calculations for performing foundation subtraction are more mind boggling than this and can be partitioned into a few classifications. The principle contrast between most strategies is the means by which the foundation model is spoken to. From easy to more complex ones:

Running Gaussian Average

For every pixel, the background is demonstrated independently as a Gaussian probability density function. The Gaussian appropriation is fitted to the n most recent pixel values and a pixel is arranged by ascertaining the probability that the most recent pixel esteem depicts an indistinguishable question from the prior pixel values did.

Mixture of Gaussians

Sometimes the part of an image that should be classified as background is not entirely static, some parts might move a little (due to wind, vibrations of the camera etc.) and should still be classified as background. To adapt to that sort of background a single valued background model is inadequate. The thought is to have distinctive Gaussian models for various conceivable background objects, if a pixel esteem is probably not going to originate from any of the diverse conveyances then it is named foreground.

Kernel density estimation (KDE)

In this method a function is constructed that gives the probability that a given pixel belongs to the distribution of background pixels. For the Gaussian running average the previous known pixel values were fitted to a Gaussian to model the distribution, in the kernel density estimator the distribution is instead constructed from a sum of kernels.

1.2 Thesis Overview

The focus of this thesis is the detection of objects in unrestricted environments monitored video cameras. The objects of intrigue are moving and in addition new static articles. The video investigation framework is not given any past information neither of the watched scene nor of the visual appearance of the articles to be recognized. The fundamental approach at the top of the priority list of the created algorithms is the location of abandoned objects out in the open spaces, which has picked up a critical consideration in the security area, since surrendered items may be regularly considered as a danger to the general population security. The final system has to provide on-line alerts to human operators. Furthermore, the detected moving objects should be provided to higher-level analysis tools in order to recognize further actions and behaviors of interest typical of surveillance systems for public spaces.

Then the background subtraction will be done by the algorithms to track object. The methods using for background subtraction are:

- An enhanced Gaussian Mixture Model (GMM) for video surveillance applications, which incorporates recent proposals for the improvement of the system performance and system convergence, and a novel heuristic for:

– Better initializing the parameters of new created modes, and

– Avoiding the emergence of over-dominating modes.

- Kernel density estimation (KDE) method a function is constructed that gives the probability that a given pixel belongs to the distribution of background pixels. For the Gaussian running average the previous known pixel values were fitted to a Gaussian to model the distribution, in the kernel density estimator the distribution is instead constructed from a sum of kernels.
- In a further video engineering is done, with background subtraction algorithm the background from one video and foreground from another video will be subtraction and we will merge them into one video. In this way, we can also do more video engineering with different videos.

CHAPTER 2
BACKGROUND SUBTRACTION
STATE OF ART

2.1 Introduction

The detection of change is a low-level vision task utilized as an initial phase in numerous computer vision applications, for example, video surveillance, low-rate video coding, human-computer connection, augmented reality or medicinal finding to say just a couple of them. Given a picture grouping, the objective is to distinguish for every frame the arrangement of pixels that are fundamentally not quite the same as the past edges. Contingent upon the application, the necessities and imperatives of the discovery calculation are distinctive. Likewise, the meaning of what is essentially unique, may rely on upon the application domain.

In the video surveillance domain, change detection has been regularly utilized as a part of request to foreground objects from the background. Foreground objects articles are the objects of automated surveillance system. The divided foreground objects are then related between frames with a specific end goal to play out a scene investigation and identify occasions of premium. In this manner, it is accepted that the background can be all around portrayed by method for a statistical model, the background model. In any case, there are some background characterstics as moving foliage or sudden brightening changes, which may make troublesome the errand of foundation displaying and upkeep. A comprehensive study of the main challenges and some principles that might be used to tackle them can be found in [18]. The segmentation of foreground objects by means of detecting the changes with reference to a background model is commonly known as background subtraction. Figure 2.1 depicts a basic schema of a general background subtraction system. The main challenges a background subtraction algorithm has to deal with are [18,19]:

- Gradual illumination changes, which are mainly experienced in outdoor environments along the different times of the day and affect the appearance of the objects in the observed scene.

- Sudden illumination changes, which are mainly experienced in indoor environments by the switching on and off of artificial light sources, and in outdoor environments by unstable weather conditions when clouds suddenly hide the sun.
- Shadows, which are mainly casted by moving objects and complicate the accurate segmentation of objects (static objects belonging to the background also cast shadows; nevertheless, these are not that problematic for the background subtraction process since they are always casted at the same position -or at slow moving positions in outdoor scenarios depending of the sun position- and can be more easily accommodated in the background model).
- Dynamic background, which are those parts of the background exhibiting different appearances because of containing some kind of moving objects as waving trees, rippling water, escalators and so on, which are not of further interest for a scene interpretation.
- Camouflage, produced by objects whose appearance is difficult to differentiate from the appearance of the background.
- Bootstrapping, which is required because of the general unfeasibility of training a background model with a completely empty scene.

Actually, in [18] the authors also pointed out some challenges which they claimed that a background maintenance system should be able to handle:

- Moved objects, which refers to the detections corresponding to background objects that have been moved.
- Sleeping person, which refers to foreground objects appearing in the scene and remaining motionless after a while.
- Walking person, which refers to objects that have been learned as part of the background and at some point in time start moving and leave the scene.

Nevertheless, these three difficulties have not been considered in this thesis as natural to the background subtraction issue, since these issues ought to be considered in agreement to the application at the top of the priority list. In fact, the point in time from which, e.g., a man nodding off is not fascinating any longer ought to be characterized by a given application and, in this way, ought not be considered as a general background upkeep issue. It is, additionally, surprising that these three issues can be likewise considered as three singularities a decent bootstrapping technique ought to handle. Finally, the foreground area gap issue, likewise specified in [18], which comprises in the unfeasibility of recognizing inside question pixels on account of shading homogeneity, has not been considered in this work as a general change discovery issue, as for the most part concerns outline differencing based methodologies.

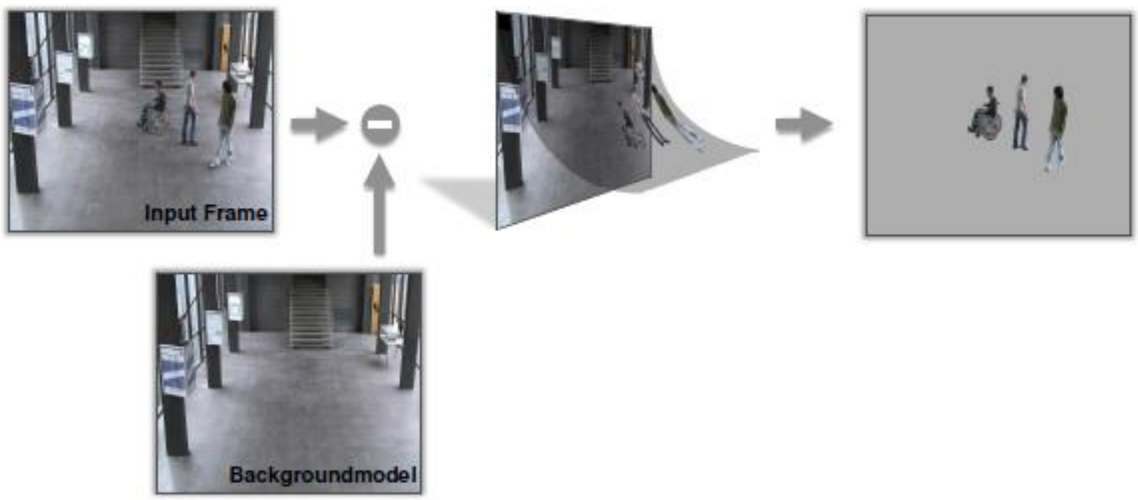


Fig 2.1 General background subtraction system[86].

2.1.1 Taxonomy

Background subtraction approaches can be divided into recursive and non-recursive. Such a taxonomy can be found in [5, 27]. Recursive approaches update the background model as new observations arrive, therefore consuming low resources in terms of computational and memory requirements. Examples of this kind of approaches can be found in [29, 28]. On the other hand, non-recursive approaches keep a buffer of the last incoming video frames to estimate the background. Therefore, non-recursive approaches have higher memory requirements. Nevertheless, since they have a copy of the most recent video frames, they can cope with some challenges as outlier rejection and fast convergence which cannot be easily handled with recursive techniques. Examples of this kind of approaches can be found in [25, 26].

NON-RECURSIVE TECHNIQUES

A Non-recursive approach utilizes sliding window idea for foundation subtraction. It cradles/buffers of pervious video casing and gauges the foundation in view of worldly variety of every pixel with in the support. For this situation, the capacity prerequisite is high, these procedures are very versatile. To fathom stockpiling issue, we can store outlines at moderate casing rate. A few strategies for the systems are portrayed underneath:

FRAME DIFFERENCING

Frame differencing uses the video frame at time, t-1, as the background model for the frame at time t, [20]. Since it uses only a single previous frame, frame differencing may not be able to identify the interior pixel of a large, uniformly colored moving object.

MEDIAN FILTERING

This is most widely used technique for background formation. The background estimate is defined to be the median at each pixel location of all he frames in the buffer, the assumption is that the pixel stays in the background for more than half of the frames in the buffer [21]. Median filtering has been extended to color by replacing the median with the Medio.

LINEAR PREDICTIVE FILTER

It computes the current background estimate by applying a linear predictive filter on the pixels in the buffer [20]. The filter coefficients are estimated at each frame time based on the sample co variances, making this technique difficult to apply in real-time.

RECURSIVE TECHNIQUE

Recursive methods don't keep up a cushion for background subtraction. Rather, they recursively overhaul a solitary background demonstrate in view of every input frame. Therefore, input outlines from removed past could affect the present background display. On the off chance that we contrasted and non-recursive method, this system requires less capacity, however any error out of sight model can proceed for a drawn out stretch of time. A few strategies for the methods are described underneath:

APPROXIMATED FILTER METHOD

This technique has been used in background modelling for urban traffic monitoring [22]. In this scheme, the running estimate of the median is incremented by one if the input pixel is larger than the estimate, and decreased by one if smaller. This estimate eventually converges to a value for which half of the input pixels are larger than and half are smaller than this value, that is, the median.

MIXTURE OF GAUSSIANS (MoG)

This method tracks multiple Gaussian distributions, MoG has enjoyed tremendous popularity since it was first proposed. This method maintains a density function for each pixel. Thus it is capable of handling multiple model background subtraction.

2.2 Relevant Approaches

2.2.1 Frame Differencing Method

In this method the difference is calculated between two frames out of which one is the current frame while the other one is the background frame to detect the presence of any moving object in the video. The equation for this is

$$|\text{frame } Ic - \text{frame } Ib| > T \qquad [23]$$

In this frame Ic is the current frame, frame Ib is the background frame and T is the threshold value.

For this the Algorithm steps are as follows: [23]

- Define the background frame and current frame from video stream.
- Calculate the gray scale converted image of those frames.
- Fix the frame dimension for further calculation of pixels.
- Calculate the difference amid pixels of the two frames and match with a defined threshold value.
- If the difference is above threshold value take it as foreground object otherwise as a background.
- Update the threshold value according to the changes in the successive frames.

The benefits of using this method is that it is fast easier to apply and performs well for static background but it needs a background not having objects otherwise they can be taken as moving object by this method.

Below are the set of research papers related to this technique:

An improved moving object detection algorithm based on frame differencing and edge detection: Zhan Chaohui [2007] state that the moving object detection and subtraction is difficult work to do. He presented an approach to detect a moving object and then subtracted it from the frame by using frame differencing method. First of all, it detects the edges of each two continuous frames and then get the difference between the two edges images. And, then it divides the edge difference image into several small blocks and decides if they are moving or steady by comparing the number of non-zero pixels

The author refers to the related work of Wan Ying [2006], Ren Mingow, Jia Zhentang [2003]

It was observed that the improved moving object detection and subtraction algorithim based on frame differencing has much greater recognition rate and higher detection speed than the several classical algorithms. This algorithm will appear individual false under more complicated background and there is still room for improvement.

Video objects extraction based on DFD between the frame and threshold segmentation: Jinwei Cui [2008] addresses the problem of complex motion and uncovered background in background segmentation, a new method was proposed based on DFD between the frames and threshold segmentation. In this method, filtering and obtained two consecutive difference between the frames and then amended the different images by “assimilation filled” to get the difference template and use the template buffer to maintain the integrity of iteration template. This algorithm doesn’t depend on a fixed background and can eliminate the uncovered background in the difference images.

The author refers the related work of Zhang Yu-Jin [1999], Jia Zhen Tang [2002], An-Ping [2006]

It was seen the result of video object extraction for single moving target video sequence is satisfactory. And it can effectively overcome the noise, the single objective of the complexity movement and the impact of background exposure in separating video object with change detection.

Object tracking using frame differencing and template matching: N. Prabhakar [2012] this author presented the object tracking and extracting system using frame differencing and template matching. The frame differencing is used frame by frame to detect a moving object in an efficient manner. The template image is used for matching purpose and generated dynamically which ensure that the change in orientation and position of object does not affect the system.

The author refers to the related work of Collins.R. [2001], V. Ramesh [2003], Yilmaz [2006]

It was observed that this method was highly effective and can be used as a surveillance tool in various applications. This method also provides better results for object extraction, which can be easily applied to a number of fields. This method can also be used to extract an object which is at a distant point. In future to improve the effectiveness more work can be done on it.

BSFD: Background Subtraction frame differencing algorithm for moving object detection and extraction: D. Stalin Alex [2014] presents the two common algorithms of moving object detection, background subtraction and frame differencing and also their comparison. The background image used to process the next frame image is generated through the super position of the current frame image. This algorithm makes the object that keep long standings, however not to be detected as a part of background.

The author refers to the related work of A. Lipton [1998], D. Gutches [2001] and Wang Ying Li [2007]

It was observed that the algorithm can detect moving object more effectively and precisely. It rectified the disadvantages of background subtraction method and frame difference method proposed a dynamic updating of background image by frame differencing method and utilises the power of the background subtraction method.

Extraction of moving objects using frame differencing, Ghost and Shadow removal. Syaimaa Solehan Mohd. Radzi [2014] this presents a technique for extracting moving objects based on temporal differencing, ghost removal using NCC, while using a non-static pan tilt zoom camera. To detect moving object in current image, the previous image frame, ft-1, is compensated with respect to the current image. This proposes a technique to remove it by using the previous image frame, ft-2. The output is then cleaned by using morphological opening operator, before shadow removal is done. Each pre-defined foreground pixels are categorized into shadow pixels or background pixels. This author refers to the work of S. Vohara [2012], D.P. Bertsekas [2004], Mc Kennel [2000]

It was observed that this method shows that the moving objects are extracted without shadows. This method can be used in real time with high computation speed and its excellent performance in detecting moving object in every frame. There are many applications which use this system, such as surveillance system in housing area, people tracking and road traffic. Future work for this project is to further improve the shadow detection with fine shape off moving objects.

Table 2.1: Summary of frame differencing

Year	Author	Title	Description
2007	Zhan Chaohui	An improved moving object detection algorithm based on frame differencing and edge detection.	Detect the problem of background subtraction in frame differencing and give the improved method to solve the problem with high detection speed and solve complicated background problem.
2008	Jinwei Cui	Video objects extraction based on DFD between the frame and threshold segmentation.	In this method author eliminates the complex motion and uncovered background and proposed new DFD method.
2012	N. Prabhakar	Object tracking using frame differencing and template matching.	Frame differencing and template matching is used to detect object and extract it effectively. This method is highly cost effective and can be used as surveillance tool in various applications.
2014	D. Stalin Alex	BSFD: Background subtraction frame differencing algorithm for moving object detection and extraction.	In this author compares the two algorithms of object subtraction. Rectified their disadvantages and proposed dynamically updated method.

2.2.2 Gaussian Mixture Model (GMM)

The Gaussian mixture model (GMM) algorithm is based on the assumption that background is more regularly visible than the foreground, and background variance is little. As a single Gaussian is not a decent model for outdoor scenes this method for background subtraction was proposed by Stauffer and Grimson [28] in which every pixel in the background is modelled as a mixture of Gaussian. Each and every pixel value is matched with current set of models to discover the match. If no match is found, the least model that is acquired is rejected and it is substituted by new Gaussian with initialization by the existing pixel value means the pixel value that don't suit into the background are taken to be background. This method requires less memory to work and gives very accurate results as well as can deal with slow lighting variation although it cannot handle multimodal background and involves rigorous computation.

Below are the set of researches related to this method:

Understanding background mixture model for background subtraction: P. Wayne Power [2002] presented the basic theory for understanding the basic model and learning by implementing Stauffer-Grimson algorithm at different parameters. It basically shows what approximations to the theory were made and how to improve the standard algorithm by redefining those approximations.

This author refers to the work of Bilmes J. [1998], Gutchess [2001], MC Ivor [2001]

It listed all the essential model parameters and typically values as well as the extension that are necessary for practical use of the algorithm. This work was providing theoretical tool with which to modify or adapt the original algorithm for better performance, higher speed and providing information needed for rapid implementation.

A Bayesian framework for Gaussian mixture background modelling: Da-Shyang Lee [2003]

It stated that background subtraction an important processing for many video applications. A Bayesian formulation of background segmentation based on Gaussian Mixture model. They show that the problem consists of two density estimation problem, one is application independent and other one is application dependent and a set of theoretically optimal solution can be derived for both. This work was tested on meeting videos and traffic videos.

This author refers to the work of A.Elgammal [2000], M. Harville [2002], C. Wren [1997]
It was showed that a set of intuitive and theoretically sound solution could be formulated in terms of density estimation problem. With this proposed algorithm the solution to these problems, the framework was applied to meeting and traffic videos segmentation. The performance over existing method validates this theory.

Improved Adaptive Gaussian Mixture model for background subtraction: Zoran Zivkovic [2004] it stated that the background subtraction is the computer task of computer vision. It is the usual pixel-level approach. In this an effective adaptive algorithm using Gaussian mixture probability density was developed recursive equation were used constantly to update the parameters and also simultaneously select the appropriate number of components for each pixel.
This author refers to the wok of C. Starffer [1999], P.J. Withagen [2002] , Z.Zivkovic [2004]
It was presented an improved GMM background subtraction scheme. This new algorithm can automatically select the needed number of components per pixel and in this way fully adapt to the observed scene. In this the processing time get reduced and segmentation also got improved.

An Improved adaptive background modelling algorithm based on Gaussian Mixture Model: Peng Suo [2008] introduced one of the best model of GMM to subtract the background scene with repetitive motion. Numerous approaches have been proposed to this problem, which differ in the type of background model, but it was one of the best. However, the large amount of computation had limited its application. Moreover, it had difficulty in segmenting slow moving objects and objects that stop for a while during moving. Based on GMM (Gaussian Mixture Model), an adaptive method was used in the algorithm to decrease the amount of the computation and an adapting method with adapting learning rate is proposed to accurately segment the objects that move slow or stop for a while.
This author refers to the work of Hou Z [2004], P. Kaer [2001] , C. Starffer [2000]
It was noticed that the comparison between the proposed algorithm and the GMM method had many differences. The segmenting results show that the proposed method had better performance than the GMM method.

Adaptive GMM approach to background subtraction for application in real time surveillance: Subra Mukherjee et. Al [2013] In this new model for real time background subtraction using a GMM (Adaptive Gaussian Mixture Model) was proposed. This new method was robout and adaptable to dynamic background, fast illumination changes repetitive motion. This also had an incorporated method for detecting shadow using the horpresert color model. This method can be used for monitoring areas where movement entry is highly restricted. So on detection of any unexpected events in the scene an alarm can be triggered and hence we can achieve a real time surveillance even in the absence of constant human monitoring.

This author refers to the frame work of W.K Wang [2009], Hao Zhou Xuejie Zhang Yun Gao Pengfei Yu [2010], Lucia Maddalena [2008]

The results of this background subtraction (AGMM) is highly effective. This method could be used to detect abandoned luggage in airport and railway stations in any place where security is prime concern. This method can be implemented so that any movement in the area can be immediately detected and an alarm can be triggered.

A novel motion object detection method based on improved frame difference and improved Gaussian Mixture Model: Yu Xiaoyang [2013] in the existing motion detection method which include background subtraction and frame difference. But it is prone to exist some holes with frame difference method and it is difficult to build a background model using background subtraction method. So previous algorithm did not achieve the ideal results. The main aim of the author is to combine frame difference method improve by motion history image with background subtraction method based on improved Gaussian mixture model to detect the motion object.

This author refers to the frame work of Lin Kai Chen [2010], Chen Ming [2012], Li Wei [2013].

It was observed that the improved frame difference was used to detect the motion object in the time domain and the improved background subtraction was used to detect the motion object in the space domain. Finally, to part were combined to obtain the complete motion object. This algorithm has processed many videos and obtain satisfactory results.

Table 2.2: Summary of Gaussian Mixture Model

Year	Author	Title	Description
2002	P. Wayne Power	Understanding background mixture model for background subtraction.	This method basically shows what approximations to the theory were made and how to improve the standard algorithm by redefining those approximations.
2003	Dar-Shyang Lee	A Bayesian framework for Gaussian Mixture Background Modelling.	A Bayesian formulation of background segmentation based on Gaussian mixture model. This also shows that the problem consists of two density estimation problems, one is application independent and other one is dependent and solution can also be derived for both.
2004	Zoran Zivkovic	Improved adaptive Gaussian mixture Model of Background subtraction.	An effective adaptive algorithm using Gaussian mixture probability density developed. The processing time get reduced and segmentation also get improved.
2008	Peng Suo	An improved adaptive background Modelling algorithm based on Gaussian Mixture Model.	In this Gaussian Mixture Model to subtract the background scene with repetitive motion. An adaptive method was used to decrease the amount of computation and accurately segment the object that move slow or stop.
2013	Subra Mukherjee*et al	An adaptive GMM approach to background subtraction for application in real time surveillance.	A new approach was proposed which is robust and adaptable to dynamic background, fast illumination changes, repetitive motion. This method can be implemented so that any movement in the area can be immediately detected and alarm can be triggered.
2013	Yu Xiaoyang	A Novel motion object detection method based on improved frame difference and improved Gaussian mixture Model.	The main aim in this is to combine frame difference method improved by motion image with background subtraction method based on improved Gaussian Mixture Model to detect the motion object. This algorithm has processed a lot videos and obtained satisfactory results.

2.2.3 Approximated Median Filter Method

McFarlane and Schofield [30] had proposed a simple recursive filter to evaluate the median of an image pixel in which the running estimate of the median is augmented by one if the input pixel is greater than the estimate and so on decremented by one if the input pixel is lesser than the estimate. The estimate ultimately converges to a value for which half of the input pixel are bigger than and half pixels are lesser than this value that is this value is the median.

In this process, the median filtering buffers the preceding N frames of the video stream. After this the background frame is computed from the median of the buffered frame and the background is subtracted from the current frame to give the foreground pixel.

The drawbacks of this technique is that it does not offer smoother results in all circumstances as it is a recursive technique it does not keeps a buffer for background estimation in its place it regularly updates a single background frame thus any input frame from a very distant past could affect the current background model. Although it means it require less memory requirements as it doesn't need to maintain a buffer.

The research papers related to this technique:

Moving vehicle segmentation in dynamic background using self-adaptive kalman background method: K.A. Ahmad [2011] this introduces the adaptive kalman filter to modeling dynamic background for background subtraction. Background subtraction method is used to identify object and famous used in moving object segmentation. This also investigate a comparision study on Gaussian subtraction method, frame differencing and approximate median method.

This author refers to the framework of Ciaran O Conaire [2006], Attila Jozsef Kun [2009], H. Kim [2008], Ya-Li How [2011].

It was observed that from kalman filter equation, we can achieve the detection of object accurately. Furthermore, the segment has been improving and the object detection more smooth.

Complex Wavelet based moving object segmentation using approximate median filter based method for video surveillance: Alok Kumar Singh Kushwahe [2014] this presented complex Wavelet based moving object segmentation using approximate median filter base method. This is capable to deal with the drawbacks such as ghosts, shadow and noise present in other spatial domain method. The performance of this method is evaluated and compared with other standard spatial domain method. Comparison is done by using relative foreground area measure, Miss-classification penalty, relative position based measure, normalized cross.

This author refers to the frame work of Y.Zhang [2006], M-Y Liu [2005], A. Khare [2008].

The obtained results and their qualitative and quantitative analysis, it can be seen that this method is performing better in comparison to other methods as well as it also capable of alleviating the problem associated with other spatial domain methods such as ghosts, clutters, noises etc.

Table 2.3: Summary of Approximated Median Filter Method

Year	Author	Title	Description
2011	K.A. Ahmad	Moving vehicle segmentation in dynamic background using self-adaptive kalman background method.	This method is new for background subtraction and also with comparison with other segmentation methods, this improves object detection and smooth segmentation.
2014	Alok Kumar Singh Kushwahe	Complex Wavelet Based Moving object segmentation using approximate median filter based method for video surveillance.	This introduces new method capable of dealing with ghosts, shadows and noise. This method is performing better in comparison to other methods as well as it also capable of alleviating the problem associated with other domain.

2.2.4 Non-parametric Model - Kernel Density Estimation

In order to cope with high-frequency variations and arbitrary distributions, non-parametric background models can be used. The probability of observing a given pixel value Xt at time t using the kernel estimator K can be non-parametrically estimated based on the pixel sample $\mathrm{X} = \{\mathrm{X}_1, \mathrm{X}_2, \ldots \ldots \ldots\ \mathrm{X}_\mathrm{N}\}$ as follows:

$$p(X_t) = \sum_{t=1}^{N} a_i\, K(X_t - X_i);$$

where α_i are weighting coefficients (usually chosen to be uniform, $\alpha_i = \frac{1}{N}$).

The probability in Equation can be efficiently computed by taking a Normal Function N (0, $\sum$) as kernel estimator, assuming independence between the different color channels, and using pre-calculated lookup tables for the kernel function given the intensity value difference ($X_t - X_i$) and the bandwidth.

The use of non-parametric background models was first proposed in [26] and [31]. In order to alleviate the high memory requirements imposed by the need of storing the whole sample set of frames considered for the density estimation, an estimation technique based on mean-shift mode finding is introduced in [32]. An approach using the balloon variable-size kernel approach, which avoids the estimation of the kernel size parameter, is proposed in [33].

However, Kernel Density Estimation (KDE)methods have a high computational cost. Moreover, in [33] it is shown that GMM seems to be a better model for simple scenes while providing a more compact representation which is suitable for further processing steps as e.g. shadow detection.

2.3 Current Trends and Conclusions

Due to its low computational load, background subtraction is presumably the most widely recognized initial phase so as to identify objects of enthusiasm for surveillance applications, particularly on account of utilizing static cameras, and has produced a broad writing. In the past segments the primary systems used to fulfill this undertaking have been introduced. These techniques have likewise been utilized in numerous other determining approaches which go for better handling a portion of the difficulties postured to the background subtraction approach. This area gives a diagram of the primary patterns saw out of background subtraction writing.

Obviously, depending on the application domain, including the characteristics of the observed scenes and computational constraints, the most suitable approach may vary. A study of various background subtraction algorithms in the context of urban traffic surveillance systems is presented in [66]. Special attention is paid to the trade-off between the obtained results and the computational complexity. The good compromise achieved by simple techniques such as adaptive median filtering for the considered domain is highlighted.

In [76], a more general selection of different methods covering a wide range of underlying mathematical approaches is presented. A categorization of the presented approaches attending to their speed, memory requirements and segmentation results is provided, aiming at facilitating the design/selection of a background subtraction approach depending on specific system requirements and capabilities. It is highlighted the acceptable accuracy provided by simple methods such as the running Gaussian average and the median filter, the high model accuracy of Gaussian mixture models and the sequential kernel approximation at the cost of higher memory and computation requirements, and the challenge posed by practical implementations to methods addressing spatial correlations.

2.3.1 Background Model Initialization

The principal undertaking to be unraveled by a background subtraction framework is the instatement of the model, regularly referred to as bootstrapping. In controlled situations, this is every now and again accomplished by forcing a preparation period during which the unfilled scene is noticeable. In any case, this methodology is not appropriate to general surveillance situations. In this manner, the background display should be introduced within the sight of moving articles. Regardless of the possibility that the utilization of straightforward methodologies, for example, a pixel-wise calculation of the mean [75] or the middle [25] esteem may suffice for a few applications, there is likewise countless, particularly those including swarms, where a more explained way to deal with foundation instatement is essential. To that point, generally some sort of spatial data is utilized. One of the most punctual methodologies in view of this guideline is displayed in [73], where the utilization of optical stream data is proposed. The primary thought is that utilizing the optical stream in the region of a pixel is conceivable to speculate if a background pixel is being impeded by a moving article (if the heading of the optical stream is towards that pixel) or if a blocked background pixel is being revealed (if the optical stream is coordinated far from that pixel). The strategy proposed in [72] comprise in processing the total of total contrasts of co-located picture block of the input frames so as to group them as moving, static closer view or static background; the background picture is figured by utilizing a worldly middle channel to join static background pieces. In [67] a strategy is proposed which comprises in isolating every information outline in patches that are bunched along the course of events keeping in mind the end goal to choose a little number of background applicants, which are then incrementally regarded to be background or not by picking at every progression the best continuation of the present foundation as indicated by visual gathering standards, subsequently considering the spatial relationships that exist inside little locales of the background picture. A later approach which likewise considers the connection of neighboring background pieces is displayed in [78], where the consolidated recurrence reaction of an applicant square and its neighborhood is the choice basis of the pieces considered as background.

2.3.2 Illumination Changes and Shadows

While steady enlightenment changes are effectively taken care of by the majority of the best in class versatile methodologies, sudden light changes and shadows threw by moving items are still a test for the majority of them. On account of worldwide illumination changes, surface and, all the more by and large, nearby based methodologies demonstrate a changeover pixel based methodologies gave that the surfaces in the watched scene are sufficiently discernable. For the instance of casted shadows, all background subtraction approaches indicate inadequacies which are typically corrected in a post-preparing step.

Sudden worldwide enlightenment changes, are typically taken care of in a spatial setting. For example, the framework proposed in [18] holds an agent set of scene background models going to various lighting conditions (a negligible set would compare to lights on and off) and picks the model that creates the least number of foreground pixels. Clearly, such an approach requires a

past learning of the vacant scene under various brightening conditions. In light of the perception that illumination changes can be better taken care of considering spatial data, the framework proposed in [68] consolidates the outcomes furnished by a GMM with spatial data gave by a disconnected spatial division of the background in a Bayesian system. A general approach which additionally exploits spatial connections is introduced in [96], where the watched scene is remedied by method for a multi-determination light revision approach keeping in mind the end goal to convey the prepared video casings to a reference luminance level. An option approach is exhibited in [77], where the foundation model is characterized by a measurable model of the light impacts, rather than the pixel powers. Besides, the probability of pixel grouping additionally melds surface connection pieces of information by misusing surface histograms prepared disconnected. Although impressive results are presented, it is assumed that the background is static and can be trained beforehand, which is a requirement that can be easily fulfilled in the scenario for which the approach is designed for, augmented reality, but not in a common
video surveillance scenario.

A survey on shadow detection approaches is presented in [97], where the different contributions reported in the literature are classified in four classes: statistical parametric, statistical non-parametric, deterministic model-based and deterministic non-model-based. Out of the evaluated approaches, the results provided by those presented in [74] and [69] are highlighted. The approach in[74] classifies pixels as foreground, background, shadowed background or highlighted background, depending on the chromaticity and brightness distortion measured by projecting the observed value into a line going through the origin of the *RGB* space and the expected value for every pixel position. The approach in [69] classifies pixels as foreground or background depending on the distance in the *HSV* color space of the observed to the expected values for every pixel position, thereby exploiting the different effect that illumination conditions have on the hue, saturation and value channels.

CHAPTER 3

METHODOLGY

This chapter serves to outline the work done during the thesis and describe how the results were obtained. The first step of the work was to get acquainted with computer vision as a field of research and the state-of-the-art in motion tracking. This was done by a literature review, especially was used to find suitable candidates for evaluation. From these two state-of-the-art trackers were chosen Gaussian Mixture model (GMM) [79] and Kernel Density estimation (KDE) [80]. In addition to these, Video engineering to be done, by background subtraction algorithm we subtract background and foreground from different videos and then we can change either background or foreground object with the new one. A system for performing the testing of algorithms was developed, written in C++ and making use of the library OpenCV for the image processing. The implementations of GMM and KDE are slight modifications of publicly available code. The code was modified to give a consistent interface for all the tracking algorithms and to make it possible to use them together with background subtraction. The results from the competition are available so it is possible to compare the results of the implementations from this thesis with that of the original algorithm authors.

An algorithm for background subtraction was implemented based on the article by Hajer Fradi [79] and Jeisung Lee [80]. One implementation of the original algorithm is available as a part of the BGSLibrary3, the implementation used in the testing in this thesis is entirely based on the written article and modified to work with a moving camera. Not all features described in [79] and [80] were implemented. The two trackers GMM and KDE were evaluated to determine whether they benefit from background subtraction. The details of evaluation is provided in next chapter.

3.1 Background subtraction

The resulting background mask from the background subtraction is used in different ways for the trackers. For GMM and KDE a new image is created from the original image by setting background pixels to black. The performance of the background subtraction is evaluated on different cases to see what impacts its performance. Two cases are constructed.

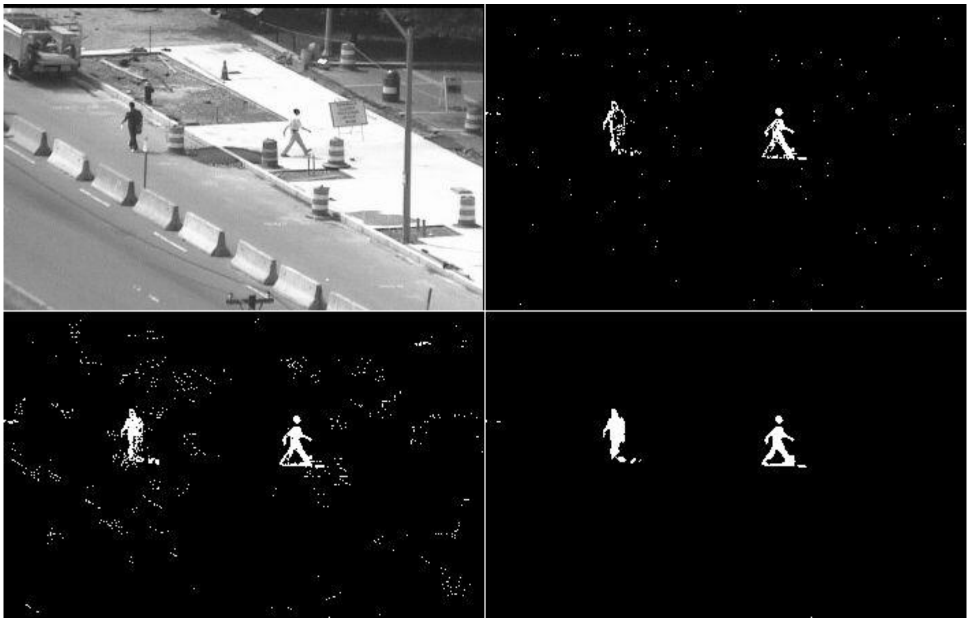


Fig 3.1 Background Subtraction Example [88]

The first is a simple sequence, taking a single large image and creating a video by sweeping over it with a smaller window. In this case, all pixels should be classified as background since there are no moving objects in a static image.

The second case is to evaluate the result of the subtraction when there is no error in the data for the camera movement. This was done by recording a sequence without moving the camera and then constructing a new video using small parts of the original sequence (moving a window over it, simulating a moving camera). By doing this we minimise vibrations and we get perfect knowledge of the per frame movement.

Finally, the background subtraction is evaluated on sequences from the camera under the conditions: only pan motions, only tilt motions, and both pan and tilt motions. For this background model, small angle rotations are assumed and the camera movement is approximated as a translation.

3.1.1 Camera parameters

To perform background subtraction when the wellspring of the frames is moving, learning about the pixelwise balance between the frames is required. One approach to get that with no earlier learning about the development is by following focuses having a place with the background starting with one casing then onto the next, by utilizing optical flow.

At the point when learning about the camera movement is accessible, some approach to relate changes in container and tilt angle to changes in pixel position in a picture is required. At the point when the adjustments in container and tilt are little, the adjustment in pixel position can be approximated to be relative to the change the position.

$$\Delta x = c_1 \,.\, \Delta p$$

$$\Delta y = c_2 \,.\, \Delta t$$

The coefficients c_1 and $c_2$can be estimated by for example using optical flow to get an estimate for the pixel movement and compare that with the change in pan and tilt. They can also be found by manually matching images with known camera position and calculate the coefficients from that. Both methods are evaluated.

3.1.2 Test procedures for evaluating the trackers with background subtraction

- Test Case 1: Tracking with background subtraction.

An arrangement of frames is gathered from a static camera. In the series, there is no less than one moving object. A moving camera is represented by building another grouping of frames where every frame is a settled size area from the relating static camera outline. The pixel position of the extricated locale is logged to simulate interpretation data from the robot. Then for each tracker:

- Initiate a bounding box on the object to be tracked.
- Track with and without background subtraction.
- If the tracking is lost, reinitialise by giving a new bounding box around the object.
- Count the number of times the tracking is lost and how many frames processed per second.

- ➢ Test Case 2: Simple tracking with background subtraction and robot. Sequences of frames are collected from the camera, in the sequence there is a moving unicoloured circle in front of a simple background. Pan and tilt positions are logged to make it possible to align the frames. Sequences:

- The camera is only panned right and left.
- The camera is only tilted up and down.
- The camera is panned and tilted in an irregular pattern.
- The moving object is moved around in an irregular pattern, the camera is manually controlled to keep the object centred in the image (as the tracker would control it).

For each tracker:

- Initiate a bounding box on the object to be tracked.
- Track with and without background subtraction.
- If the tracking is lost, reinitialise by giving a new bounding box around the object.
- Count the number of times the tracking is lost and how many frames processed per second.

- ➢ Test Case 3: Tracking with background subtraction and robot
- Same procedure as scenario 2 but with a more complicated background with clutter.

3.2 Gaussian Mixture Model Algorithm

To account for complex backgrounds containing more than one Gaussian distribution, [28] models each pixel as a mixture of K Gaussians corresponding to either background or foreground. The probability of the occurrence of a current pixel is [28]:

$$P(I_{p,t}) = \sum_{i=1}^{K} w_{i,p,t} * \eta(I_{p,t}\ ;\ \mu_{i,p,t})$$

where η($\mu_{i,p,t}, \sum_{i,p,t}$) is the i^{th} background Gaussian model and $\omega_{i,p,t}$ its weight. Pixel values that do not fit the background distributions are considered as foreground until there is sufficient and consistent evidence to initiate a new Gaussian to support them. The background Gaussians can be determined in terms of its persistence and the variance which can be measured by ω/σ. This value increases both as a distribution gains more confidence and more persistent. After ordering the Gaussians by ω/σ, the first B distributions are chosen as the background model, where [28]

$$B_{p,t} = \arg\min_{b}(\sum_{i=1}^{b} \omega_{i,p,t} > T)$$

where T is a measure of the minimum portion of the data that should belong to background. Thus $I_{p,t}$ is labeled as background if it is standard deviation of a background Gaussian model. GMM has gained vast popularity [83, 85, 82, 84]. Yet [81] points out that it fails to achieve sensitive detection in the case where the background has very high frequency variations such as waving water or shaking tree leaves, i.e., background having fast variations cannot be accurately modelled with just a few Gaussians. Another important point is its ability to adapt to sudden change in the background which depends on the learning rate. Low learning rate is suitable for long-term change but it has a poor adaptivity to sudden change. High learning rate can adapt to changes quickly, but slowly moving objects can be easily incorporated into background.

Fig 3.2 Example of Gaussian mixture Model [89]

3.3 Kernel Density Estimation Algorithm

The kernel density estimation (KDE) method, a non-parametric approach that can effectively adapt to a dynamic background. In each pixel, the KDE is calculated by the following equation at time index *t [80]*:

$$p(x) = \frac{1}{n}\sum_{i=1}^{n} K(x - x_t)$$

where n is the number of total observed frames and x_t is the observed value at time index t. p(x) is an average of normal densities centered at the sample x. The kernel function K(x) should satisfy the following conditions: $\int K(x)dx = 1, \quad \int xK(x)dx = 0, and\ K(x) > 0$. Typically, the normal distribution N (0,1) is used as the kernel function. In research conducted by Park et al., many frames were collected before estimating the Gaussian background model and thus, a large amount of memory space was required. To overcome this drawback, we modify the original KDE method and propose a scheme that uses the first frame to initialize the KDE background model. In the first frame, most of the pixels represent background, and there are foregrounds in some other pixels. Even if we used the first frame to initialize background model, foreground information will be reduced and remain only background information by updating process because background values are more frequent than foreground values at the pixel level. The KDE Gaussian model is subsequently updated at every frame by controlling the learning rate according to the situation. The probability $p_t(x)$ is based on each pixel and may be expressed as [80]:

$$p_t(x) = \hat{p}_{t-1}(x) + \frac{1}{G_t\sqrt{2\pi\sigma^2}} \exp\left[\frac{1}{2}\left(\frac{x - x_t}{\sigma}\right)^2\right]$$

Each pixel has a probability model. The probability obtained by the KDE method is added to the prior probability density at every frame. In second equation, G_t is used as the learning rate at time t and can be changed depending on factors such as time and illumination changes. Since the probability should satisfy $\int p_t(x)dx = 1, p_t(x)$ is normalized as follows [80]:

$$\hat{p}_t(x) = p_t(x) \,/ \sum_{x=0}^{N} p_t\,(x)$$

where $p_t(x)$ is a normal density at the sample x and at time index t. $\hat{p}_t(x)$ is a normalized normal density and N is the total number of samples. A new probability background model is obtained through the above process. This updating method improves memory effectiveness because it does not require many images to be saved to initialize the probability background model. The updating method automatically reduces the probability of unimportant backgrounds that do not appear over a long period by adding an additional probability and performing a normalization step. For example, when a car parked for a long period moves or disappears, the proposed method continually updates the environment. Consequently, new background information appears and the prior unimportant background probability associated with the car is automatically lowered by updating the background model. We used G_t as a parameter to control the learning rate. If G_t is increased, new information is slowly learned and prior information slowly disappears. If G_t is decreased, the algorithm quickly adapts to the environment and quickly deletes old information. In the initial stage, the background model should quickly adapt to the new environment and, as time elapses, the background should have a stable updating process. For this reason, G_t was used as a sigmoid function which can expressed as follows [80]:

$$G_t = Gain * \frac{2}{1 + \exp(-(cnt - \beta)/\lambda}$$

A few of the problems associated with the non-parametric kernel density estimation approach are the undesirably long processing time and the large memory requirement. We can reduce the complexity and memory requirement using histogram approximation. The Gaussian probability and an example of histogram approximation. B_d is the width of the histograms along dimension d, C_k is the center of each histogram, and k is the histogram number? The parameter $B_d$can be calculated according to the following equation [80]:

$$B_d = \frac{\max(x^d) - \min(x^d)}{N_d} \qquad\qquad d = 1,2,3$$

where N_d represents the number of bins for each dimension d and x^d is the value of a pixel in the d dimension.

Fig 3.3 Example of KDE [90]

3.3.1 Shadow Detection

To remove the shadows of moving objects, we applied a moving cast shadow detection algorithm [70] that proved to be quite accurate and suitable for eliminating shadows. The basic idea is that a cast shadow darkens the background, while the color of the background itself is not changed. Using this principle, we can express the removing shadow algorithm as follows [80]:

$$[\rho \leq \frac{x^v}{Bg^v} \leq \delta] \wedge (|x^s - Bg^s| \leq \tau_s) \wedge (|x^h - Bg^h| \leq \tau_h)$$

where Bg^h, Bg^s, and Bg^v represent the hue, saturation, and illumination components, respectively, of the background pixels with background values that are closest to the input image among background histogram models. x^v, x^s, and x^h represent the hue, saturation, and illumination components of the input video pixels.

3.3.2. Adaptation for a Sudden Illumination Change

If the background itself is significantly changed (e.g., suddenly brightened or darkened), fast adaptation is required. We can obtain this effect by initializing the *cnt* value. If the value of *cnt* is initialized, G_t is also initialized and the speed of adaptation for the background increases [80]:

$$\begin{Bmatrix} Mv_t = (G_t - 1 * \frac{Mv_{t-1}}{G_t} + mean_{\forall_{i,j}}(Dist_v(i,j))/G_t \\ if\left(\left|mean_{\forall_{i,j}}\left(Dist_v(i,j)\right) - Mv_t\right| > T_v\right)\ then, cnt = {}^{\beta}/_{2} \end{Bmatrix}$$

T_v is a threshold to initialize *cnt*; it is set to 30 in our experiments. $Dist_v(i, j)$ is a illumination value of current input image at the (i, j) pixel. Mv_t is an moving average value of $mean_{\forall i,j}(Dist_v(i, j))$.S

CHAPTER 4
EXPERIMENTAL RESULTS
AND DISCUSSION

In this chapter, we conduct a set of experiments. We use GMM and KDE to model the background, for the subtraction [79,80] and for video engineering, for updating background [25,75,73,72]or foreground object. Yet as we discussed in chapter 2 and 3 through a state-of-art and detailed algorithms.

Different competing video sequences with resolution of 240 × 320 at 30 frames per second were used to analyze the performance of the background subtraction approaches in different environments.

The detection results are presented qualitatively and quantitatively. The parameters for each algorithm were determined experimentally. For each sequence, several representative frames, the ground truth and detection results produced by each algorithm are presented. The detection results are shown as black and white images where white pixels represent foreground objects while black pixels represent background. The performance of each approach is also evaluated quantitatively using a) the traditional pixel wise evaluation metrics (precision, recall, F-measure) which are used commonly in evaluating background subtraction approaches and b) the component-based evaluation metrics which are designed from the perspective of object detection, here we use the correct detection rate, miss detection rate and false alarm rate defined in [87].

4.1 Evaluation Metrics

In this thesis, we use two types of measurements to evaluate the performance of different approaches, one defined in pixel-level, the other in component-level.

The first type of evaluation metrics defined in pixel level is the most direct measure which is often used often to evaluate the performance of background subtraction approaches, including precision, recall and F-measure. They are defined as follows:

$$Precision = \frac{\#True\ Positives}{\#true\ Positives + \#False\ Postives}$$

$$Recall = \frac{\#True\ Postives}{\#True\ Positives + \#False\ Negatives}$$

$$F - measure = 2\ .\frac{Precision\ .recall}{Precision + Recall}$$

These evaluation metrics measure the accuracy of the approach at the pixel level, however, in some cases, people are not interested in the detection of point targets but object regions instead. Thus, we also use the object-based evaluation metrics proposed in [87].

To be more specific, we consider three cases mentioned in [87] which are shown as follows:

• **Correct Detection (CD) or 1-1 match**: the detected region corresponds to one and only one ground truth region.

• **False Alarm (FA)**: the detected region has no correspondence in the ground truth.

• **Detection Failure (DF)**: the ground truth region is not detected.

According to the definitions, we need to determine the correspondence of the foreground region in the detection result and in the ground truth, i.e., whether the foreground region in the ground truth is matched with the segmentation. Based on the correspondences, we can evaluate a selected approach in terms of the correct detection rate, the false alarm rate and the detection failure rate.

4.2 Experimental Results

In this chapter, experimental results comparing different approaches of background subtraction algorithms are presented. Experiments are conducted on different sequences, which demonstrate that our approach outperforms among these algorithms and it is robust to the outliers from inaccurate motion estimates, and pixel misalignment when registering consecutive images. The result of comparing the appearance-based approaches with that of incorporating motion and appearance demonstrates that, motion can provide higher discriminative power than using appearance cue alone, which can improve the robustness to the outliers from image registration, yet modelling motion and appearance cues jointly is vulnerable towards these outliers from either cue, since these outliers may be introduced into the joint kernel function, which will deteriorate its accuracy. Evaluating marginal probabilities is useful to deal with the outliers and provides higher precision, yet the recall may be much lower since it may be overly conservative.

4.2.1Background Subtraction results

Table 4.1: Qualitative comparison of video Sequence 1

$\boldsymbol{i^{th}}$	$\mathbf{155^{st}}$	$\mathbf{221^{st}}$	$\mathbf{591^{st}}$	$\mathbf{811^{st}}$
IMG				
GMM				
KDE				

Table 4.2: Qualitative comparison of video sequence 2

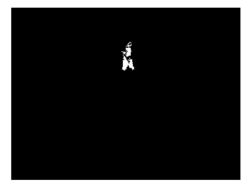
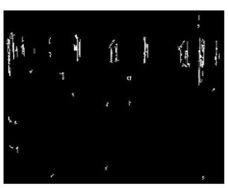
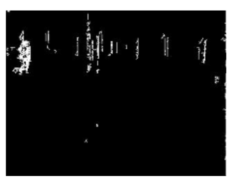
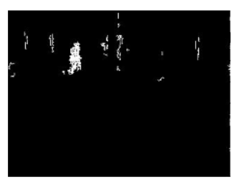
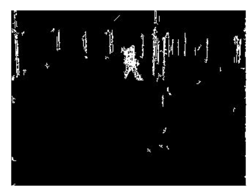
ith	**125^{th}**	**237^{th}**	**272^{nd}**	**305^{th}**
IMG				
GMM				
KDE				

Table 4.3: Qualitative comparison on indoor sequence 3.

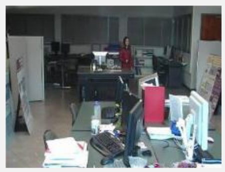
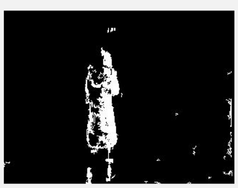
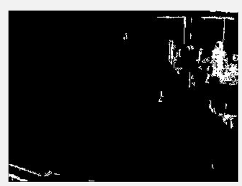
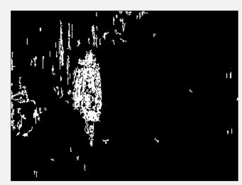
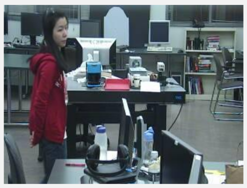
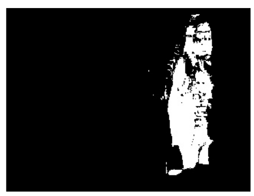
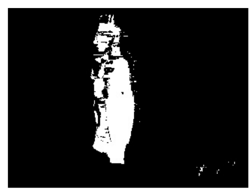
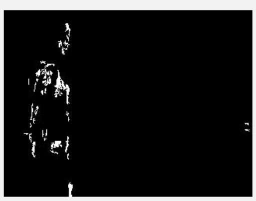
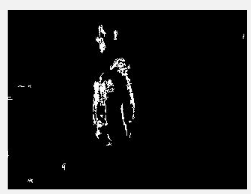
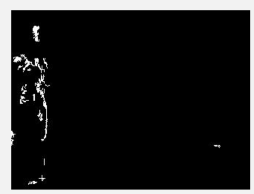
Ith	***62^{nd}***	***97^{th}***	***191^{st}***	***345^{th}***
IMG				
GMM				
KDE				

Table 4.4: Qualitative comparison on indoor sequence 4

Ith	*61st*	*78th*	*98th*	*130th*
IMG				
GMM				
KDE				

4.2.2 Video Engineering Results

(a) (b) (c)

Fig4.1 (a) Containing foreground object as input (b) representing the stationary or background of scene (c) representing output with the new foreground object.

Fig 4.2 Result for background updating

4.3 Discussion:

In this chapter, experimental results comparing two approaches of background extraction are presented. Experiments are conducted on different video sequences, which demonstrate that GMM approach outperforms the other algorithm and is robust to outliers coming from inaccurate motion estimation and pixel misalignment, when registering consecutive images. The result of comparing the appearance-based approaches with that of incorporating motion and appearance demonstrates that, motion can provide higher discriminative power than using appearance cue alone, which can improve the robustness to the outliers from the image registration, yet modelling motion and appearance cues jointly is vulnerable towards these outliers from either cue, since these outliers may be introduced into the joint kernel function, which will deteriorate its accuracy.

Moreover, results of video engineering is also shown in which we did video editing, for instance we have two videos with which we changed background of one video with other video and in one video we changed the foreground objects with the new objects with same background. Video engineering was mostly done by manual method but in our case, we did it automatically. We fed the videos to the algorithm and it automatically replaces the objects. This method could be useful to the movie industry and also helpful in video surveillance.

CHAPTER 5
CONCLUSION

This thesis deals with the detection of objects of arbitrary visual appearance in surveillance video data. In particular, the objects of interest were of two different natures: moving objects, which pass by through the observed scene, and static objects, which are added or removed from the scene. Moving objects should be provided to higher-level analysis layers for action and behavior recognition. Static objects should provide on-line alerts to human operators in real-time.
The absence of appearance models (and the unfeasibility to build them) and the immobility of the static objects has led to the use of background subtraction as the low-level processing tool. A thorough review of state-of-the-art background subtraction methods has been provided, thereby highlighting the main problems faced by this technique and how these problems have been approached in the extensive literature.

Gaussian Mixture Models (GMM) and Kernel Density model have been chosen as the underlying background models. In this thesis, background subtraction was done by both methods and we have found that the GMM is the better approach for background subtraction and with its help, we were able to perform some video engineering as well. We replaced the foreground object of a video with another object, extracted from another video.

Video engineering is usually a manual process, but with the help of this method it was possible to do it automatically. The only manual part in our approach is to provide the input videos to the method. Video engineering could also be useful for video editing, like making changes to videos.

REFRENCES

1. Appiah, K. andHunter, A. (2005). A single-chip fpga implementation of real-time adaptive background model. In Proceedings of the IEEE International Conference on Field-Programmable Technology, pages 95–102
.
2. Azzari, P. and Bevilacqua, A. (2006). Joint spatial and tonal mosaic alignment for motion detection with ptz camera. In Proceedings of the ICIAR.

3. Babaguchi, N., Fujimoto, Y., Yamazawa, K., and Yokoya, N. (2002). A system for visualization and summarization of omnidirectional surveillance video. In Proceedings of the 8th International Workshop on Multimedia Information Systems (MIS2002), pages 18–27, Tempe AZ.

4. Baker, S., Scharstein, D., Lewis, J., Roth, S., Black,M. J., and Szeliski, R. (2011). A database and evaluation methodology for optical flow. International Journal of Computer Vision, pages 1–31.

5. Baltieri, D., Vezzani, R., and Cucchiara, R. (2010). Fast background initialization with recursive hadamard transform. In Proceedings of the IEEE International Conference on Advanced Video and Signal Based Surveillance, pages 165–171.

6. Barnich, O. and Van Droogenbroeck, M. (2011). Vibe: A universal background subtraction algorithm for video sequences. IEEE Transactions on Image Processing, 20(6):1709 –1724.

7. http://electronicimaging.spiedigitallibrary.org/article.aspx?articleid=2457406[Last accessed on 15/01/2017]

8. W. Hu, T. Tan, L.Wang, and S. Maybank,\A survey on visual surveillance of object motion and behaviors," Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, vol. 34, no. 3, pp. 334{352, 2004.

9. Basu, M. (2002). Gaussian-based edge-detection methods-a survey. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 32(3):252–260.

10. Bay, H., Ess, A., Tuytelaars, T., and Gool, L. V. (2008). Surf: Speeded up robust features. Computer Vision and Image Understanding (CVIU), 110:346–359.

11. Bayona, Á., SanMiguel, J. C., andMartínez, J.M. (2009). Comparative evaluation of stationary foreground object detection algorithms based on background subtraction techniques. In Proceedings of the IEEE International Conference on Advanced Video and Signal Based
Surveillance, pages 25–30.

12. Beauchemin, S. S. and Barron, J. L. (1995). The computation of optical flow. ACM Computing Surveys, 27(3):433–467.

13. Benezeth, Y., Jodoin, P.-M., Emile, B., Laurent, H., and Rosenberger, C. (2010). Comparative study of background subtraction algorithms. J. Electronic Imaging, 19.

14. Bevilacqua, A., Di Stefano, L., and Azzari, P. (2005). An effective real-time mosaicing algorithm apt to detectmotion through background subtraction using a ptz camera. In Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, pages 511 – 516.

15. Bhat, K., Saptharishi,M., and Khosla, P. (2000). Motion detection and segmentation using image mosaics. InMultimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on, volume 3, pages 1577 –1580 vol.3.

16. Bilmes, J. A. (1998). A gentle tutorial on the EM algorithm and its application to parameter estimation for gaussian mixture and hidden markov models. Technical report, International Computer Science Institute, Berkeley CA.

17. Liebowitz, D., Criminisi, A., and Zisserman, A. (1999). Creating architectural models from images. In Annual Conference of the European Association for Computer Graphics (Eurographics), volume 18, pages 39–50.

18. Toyama, K., Krumm, J., Brumitt, B., and Meyers, B. (1999). Wallflower: Principles and Practice of Background Maintenance. In Proceedings of the IEEE International Conference on Computer Vision, volume 1, page 255, Los Alamitos, CA, USA. IEEE Computer Society.

19. Brutzer, S., Hoferlin, B., and Heidemann, G. (2011). Evaluation of Background Subtraction Techniques for Video Surveillance. In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), pages 1937–1944, Colorado Spring, USA.

20. U.Chandrasekhar, Tapan Kumar Das, "A survey technique for background subtraction and traffic analysis on Surveillance Video", 2011.

21. Ye Song, Na Fu, Xiaoping Li, Qiongxiu Liu, "Fast moving object detection using improved Gaussian Mixture method" ,2014.

22. R. Cucchiara, M. Piccardi, "Detecting moving objects, ghost and shadows in video stream ", 2013.

23. P.Remagnino et. Al, "An integrated traffic and pedestrian model-based vision system", 1997.

24. Baltieri, D., Vezzani, R., and Cucchiara, R. (2010). Fast background initialization with recursive hadamard transform. In Proceedings of the IEEE International Conference on Advanced Video and Signal Based Surveillance, pages 165–171.

25. Cutler, R. and Davis, L. (1998). View-based detection and analysis of periodic motion. In Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on, volume 1, pages 495 –500 vol.1.

26. Elgammal, A., Harwood, D., and Davis, L. S. (2000). Non-parametric model for background subtraction. In Proceedings of the 6th European Conference on Computer Vision.

27. Parks, D. H. and Fels, S. (2008). Evaluation of background subtraction algorithms with postprocessing. In Proceedings of the IEEE International Conference on Advanced Video and Signal Based Surveillance, pages 192–199.

28. Stauffer, C. and Grimson, W. (1999). Adaptive Background Mixture Models for Real-time Tracking. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, volume 2, pages 246–252, Fort Collins, CO, USA.

29. Wren, C., Azarbayejani, A., Darrell, T., and Pentland, A. (1997). Pfinder: Real-time tracking of the human body. IEEE Transactions on Pattern Analysis andMachine Intelligence, 19:780– 785.

30. N. McFarlane and C.Scofield "Segmentation and tracking of piglets in images", Machine vision and application, 1995.

31. Elgammal, A., Harwood, D., and Davis, L. S. (2000). Non-parametric model for background subtraction. In Proceedings of the 6th European Conference on Computer Vision.

32. Han, B., Comaniciu, D., and Davis, L. (2004). Sequential kernel density approximation through mode propagation: Applications to background modelling. In Proceedings of the Asian Conference on Computer Vision.

33. Zivkovic, Z. and van der Heijden, F. (2006). Efficient adaptive density estimation per image pixel for the task of background subtraction. Pattern Recognition Letters, 27:773–780.

34. S. Indupalli, B. Boufama, "A Novel Clustering-Based Method for Adaptive Background Segmentation", 2006.

35. Zhihai Sun, Shan-an Zhu, "Real-Time and Automatic Segmentation Technique for Multiple Moving Object in Video Sequence", 2007.

36. Yong Dong Zhang, Jia Li Zheng, "Hierarchical Threshold Technique Oriented to Video Background Segmentation", 2007.

37. Jinwei Cui, Bing Wang, "Video Objects Extraction Based On DFD Between the Frames and Threshold Segmentation", 2008.

38. Lakis Christodoulou, Takis Kasparis Oge Marques.an, "Advanced Statistical and Adaptive Threshold Techniques for Moving Object Detection and Segmentation", 2011.

39. Dingming Liu, Jieyu Zhao, "Spatio-temporal Video Object Segmentation Using Moving Detection and Graph Cut Methods" ,2011.

40. U.Chandrasekhar, Tapan Kumar Das, "A survey technique for background subtraction and traffic analysis on Surveillance Video", 2011.

41. R.Revathi, M.Hemalatha, "Detecting Objects in Video Frames using Optical Flow Techniques " ,2012.

42. Yu Xiaoyang, Yu Yang, Yu Shuchun, Song Yang, "A Novel Motion Object Detection Method Based on Improved Frame Difference and Improved Gaussian Mixture Model", 2013.

43. Adaptive Threshold based Segmentation for Video Object Tracking. [2014] - Deepak Gambhir, Meenu Manchanda.

44. Guoheng Huang, Chi-Man Pun and Cong lin, "Video Object Tracking Using Interactive Segmentation and Superpixel based Gaussian Kernel", 2015.

45. "Fast Approximate Matching of Videos from Hand-Held Cameras for Robust Background Subtraction" ,2015.

46. Stefan Becker, Norbert Scherer-Negenborn, "An evaluation of background subtraction algorithm on fused infrared-visible video streams", 2015.

47. ” Real-Time and Robust Compressive Background Subtraction for Embedded Camera Networks “, 2016.

48. E. Komagal, Arthy Vinodhini, Archana Bricilla, “Real time Background Subtraction Techniques for Detection of Moving Objects in Video Surveillance System”, Velammal college of Engg. And Tech., Madurai-625009, Tamil Nadu, India.

49. Mr. D. W. Chinchkhede1 & Mr. N. J. Uke, “Image Segmentation in video Sequence using modified Background Subtrsaction”, SVC Polytechnic, Pune University, Pune, India.

50. MA Zhongli, WEN Jie, HAO Liangliang, WANG xiang, “Targets Image Segmentation Based on Adaptive Mean-Shift Algorithm for Surface Moving Platform”, 2014.

51. K.Srinivasan, K.Porkumaran and G.Sainarayanan, “Improved background subtraction techniques for security in Video applications”, 2009.

52. Zhan Chaohui, “An improved moving object detection algorithm based on frame differencing and edge detection:”, 2007.

53. Jinwei Cui, “Video objects extraction based on DFD between the frame and threshold segmentation”, 2008.

54. N. Prabhakar, “Object tracking using frame differencing and template matching”, 2012.

55. D. Stalin Alex, “BSFD: Background Subtraction frame differencing algorithm for moving object detection and extraction:”, 2014.

56. Syaimaa Solehan Mohd. Radzi, “Extraction of moving objects using frame differencing, Ghost and Shadow removal”, 2014]

57. P. Wayne Power, “Understanding background mixture model for background subtraction”, 2002.

58. Da-Shyang Lee, “A Bayesian framework for Gaussian mixture background modelling”, 2003.

59. Zoran Zivkovic, “Improved Adaptive Gaussian Mixture model for background subtraction” ,2004.

60. Peng Suo, “An Improved adaptive background modelling algorithm based on Gaussian Mixture Model”, 2008.

61. Subra Mukherjee et. Al, “An Adaptive GMM approach to background subtraction for application in real time surveillance”, 2013.

62. Yu Xiaoyang, “A novel motion object detection method based on improved frame difference and improved Gaussian Mixture Model”, 2013.

63. Ye Song, “Fast moving object detection using improved Gaussian Mixture Model”, 2014.

64. K.A. Ahmad, “Moving vehicle segmentation in dynamic background using self-adaptive kalman background method” ,2011.

65. Alok Kumar Singh Kushwahe, “Complex Wavelet based moving object segmentation using approximate median filter based method for video surveillance”, 2014.

66. Cheung, S.-C. S. and Kamath, C. (2004). Robust techniques for background subtraction in urban traffic video. In Proceedings of the IS&T/SPIE’s Symposium on Electronic Imaging, SannJose, CA, United States.

67. Colombari, A., Fusiello, A., and Murino, V. (2006). Background initialization in cluttered sequences. In Proceedings of the 2006 Conference on Computer Vision and Pattern Recognitionn Workshop, CVPRW’06, pages 197–, Washington, DC, USA. IEEE Computer Society.

68. Cristani, M., Bicego, M., andMurino, V. (2002). Integrated region- and pixel-based approach to background modelling. In Proceedings of the IEEE Workshop on Motion and Video Computing, pages 3–8, Orlando, FL.

69. Cucchiara, R., Grana, C., Neri, G., Piccardi, M., and Prati, A. (2002). The sakbot system for moving object detection and tracking. In Video-Based Surveillance Systems, pages 145–157. Springer US.

70. Cucchiara, R., Grana, C., Piccardi, M., and Prati, A. (2003). Detecting moving objects, ghosts, and shadows in video streams. IEEE Transactions on Pattern Analysis andMachine Intelligence, 25(10):1337–1342.

71. Cutler, R. and Davis, L. (1998). View-based detection and analysis of periodic motion. In Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on, volume 1, pages 495 –500 vol.1.

72. Farin, D., deWith, P., and Effelsberg,W. (2003). Robust background estimation for complex video sequences. In Proceedings of the IEEE International Confeference in Image Processing (ICIP), pages 145–148, Barcelona.

73. Gutchess, D., Trajkovi´c,M., Cohen-Solal, E., Lyons, D., and Jain, A. K. (2001). A background model initialization algorithmfor video surveillance. In in proceedings of the IEEE International Conference on Computer Vision, pages 733–740.

74. Horprasert, T.,Harwood, D., andDavis, L. S. (1999). A Statistical Approach for Real-time Robust Background Subtraction and Shadow Detection. In Proceedings of the IEEE Conference ICCV, volume 99, pages 1–19.

75. Koller, D., Weber, J., Huang, T., Malik, J., Ogasawara, G., Rao, B., and Russell, S. (1994). Towards robust automatic traffic scene analysis in real-time. In Proceedings of the 33rd IEEE Conference on Decision and Control, volume 4, pages 3776–3781 vol.4.

76. Piccardi, M. (2004). Background subtraction techniques: a review. In Proceedings IEEE International Conference on Systems,Man and Cybernetics, volume 4, pages 3099–3104.

77. Pilet, J., Strecha, C., and Fua, P. (2008). Making background subtraction robust to sudden illumination changes. In Proceedings of the ECCV.

78. Reddy, V., Sanderson, C., and Lovell, B. C. (2011). A low-complexity algorithm for static background estimation from cluttered image sequences in surveillance contexts. EURASIP Journal on Image and Video Processing.

79. Hajer Fradi and Jean-Luc Dugelay, Robust Foreground Segmentation Using Improved Gaussian Mixture Model and Optical Flow.

80. Jeisung Lee and Mignon Park*, (2012), An Adaptive Background Subtraction Method Based on Kernel Density Estimation

81. Ahmed Elgammal, David Harwood, and Larry Davis. Non-parametric model for background subtraction. 6th European Conference on Computer Vision, pages 751–767, 2000.

82. Dar-Shyang Lee. Effective gaussian mixture learning for video background subtraction. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(5):827–832, 2005.

83. Horng Horng Lin, Tyng Luh Liu, and Jen Hui Chuang. Learning a scene background model via classification. IEEE Transactions on Signal Processing, 57(5):1641–1654, 2009.

84. Yang Wang Kia-Fock Loe and Jian-Kang Wu. A dynamic conditional random field model for foreground and shadow segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(2):279–289, 2006.

85. Richard J. Radke, Srinivas Andra, Omar Al-Kofahi, and Badrinath Roysam. Image change detection algorithms: A systematic survey. IEEE Transactions on Image Processing, 14(3):294–307, 2005.

86. Background Subtraction for the Detection of Moving and Static Objects in Video Surveillance,2014.

87. Jacinto Nascimento and Jorge Marques. Performance evaluation of object detection algorithms for video surveillance. IEEE Transactions on Multimedia, pages 761–774, 2006.

88. M.J. McHugh, J. Konrad, V. Saligrama, D. Castanon Foreground-Adaptive Background Subtraction Air Force Office of Sponsored Research (SBIR), College of Engineering Catalyst Award. (2007-2008)

89. http://docs.opencv.org/3.1.0/db/d5c/tutorial_py_bg_subtraction.html. [last accessed on 15/01/2017]

90. Non-parametric Background Subtraction, Ahmed Elgammal - David Harwood - Larry Davis Computer Vision Laboratory, University of Maryland, College Park, MD, 20742, USA

91. Szeliski, R. (2010). Computer Vision: Algorithms and Applications. Springer.

92. Jain, R. and Nagel, H.-H. (1979). On the analysis of accumulative difference pictures from image sequences of real world scenes. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-1(2):206–214.

93. Broida, T. J. and Chellappa, R. (1986). Estimation of object motion parameters from noisy images. Pattern Analysis and Machine Intelligence, IEEE Transactions on, PAMI-8(1):90–99.

94. Tanizaki, H. (1987). Non-gaussian state-space modeling of nonstationary time series. J. Amer. Statist. Assoc., 82:1032–1063.

95. Yilmaz, A., Javed, O., and Shah, M. (2006). Object tracking: A survey. ACM Computer Surv ., 38.

96. Suau, X., Casas, J. R., and Ruiz-Hidalgo, J. (2009). Multi-resolution illumination compensation for foreground extraction. In ICIP'09: Proceedings of the 16th IEEE International Conference on Image Processing, pages 3189–3192, Piscataway, NJ, USA. IEEE Press.

97. Prati, A.,Mikic, I., Trivedi, M., and Cucchiara, R. (2003). Detectingmoving shadows: algorithms and evaluation. Pattern Analysis andMachine Intelligence, IEEE Transactions on, 25(7):918 –923.

VITA AUCTORIS

NAME:	Shanpreet Kaur
PLACE OF BIRTH	Punjab, India
YEAR OF BIRTH	1991
EDUCATION	Guru Nanak Public School, India
	2008-2010
	Chandigarh Group of Colleges, India
	2010-2014 B.Tech.
	University of Windsor, Windsor, Ontario
	2015-2016 M.Sc.